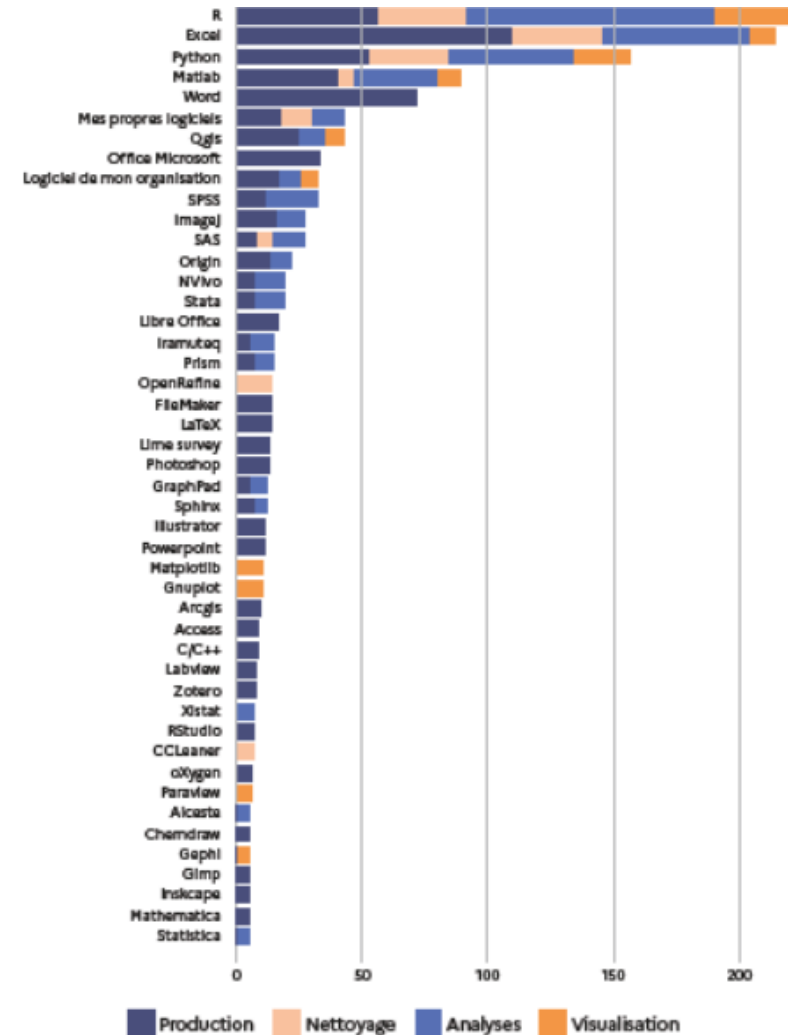


Ulrike LUCKE

University of Potsdam, Germany

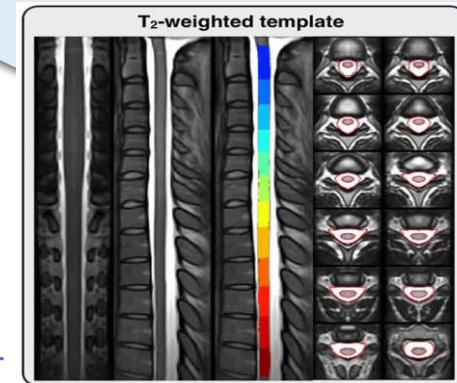
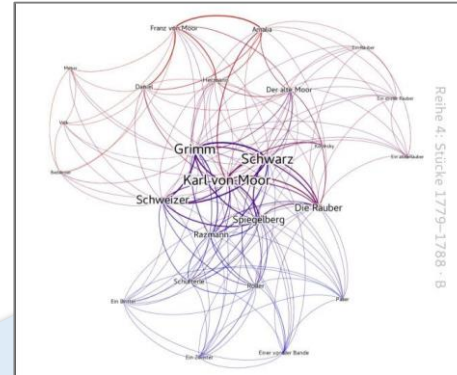
The role of Infrastructure for Software in Open Science

Software in Research

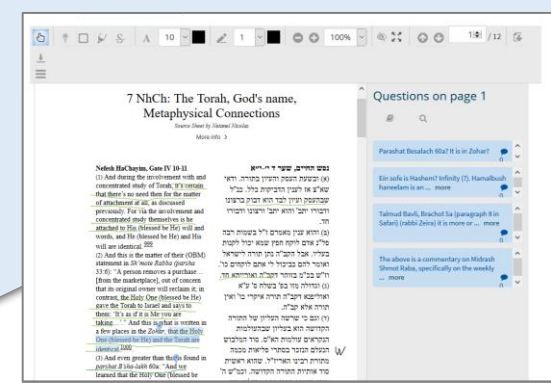
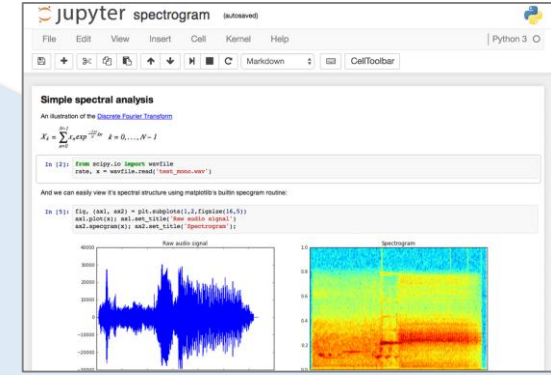


<https://tinyurl.com/opensciencepractices>

software
edition taberna kritika
Spinal Cord Toolbox



general-purpose
special-purpose



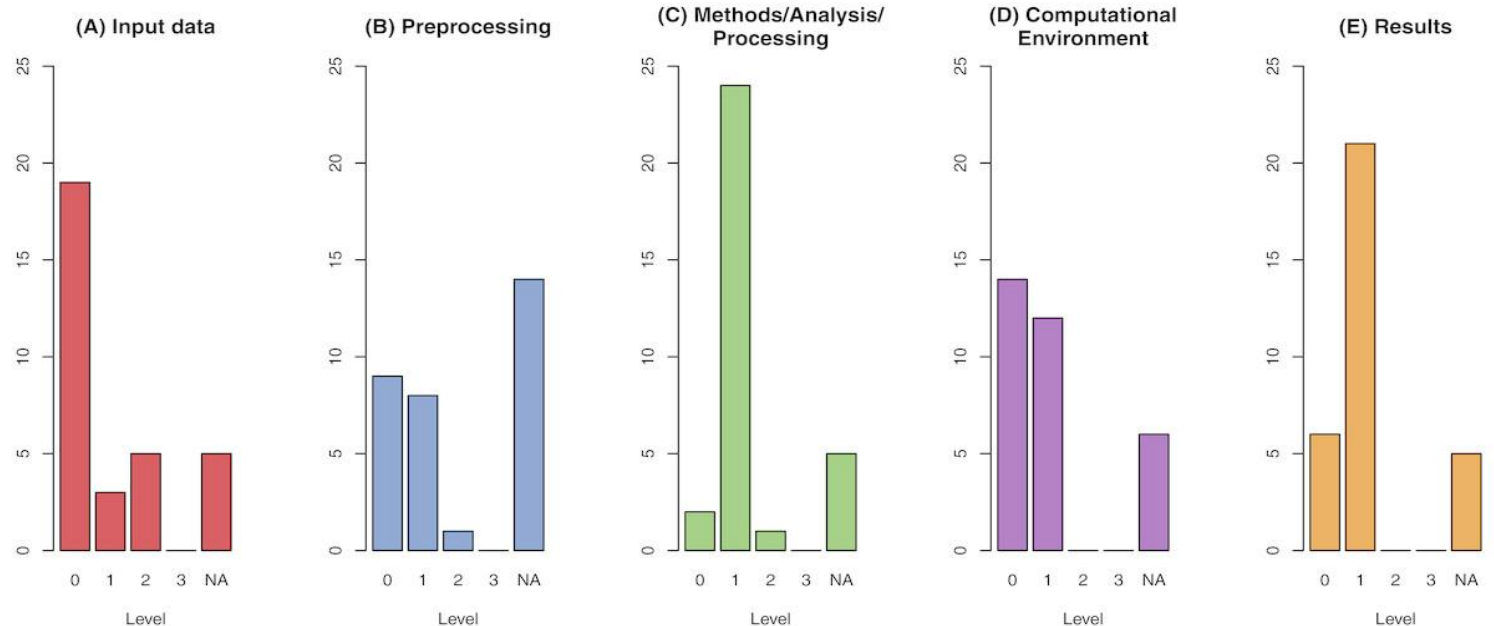
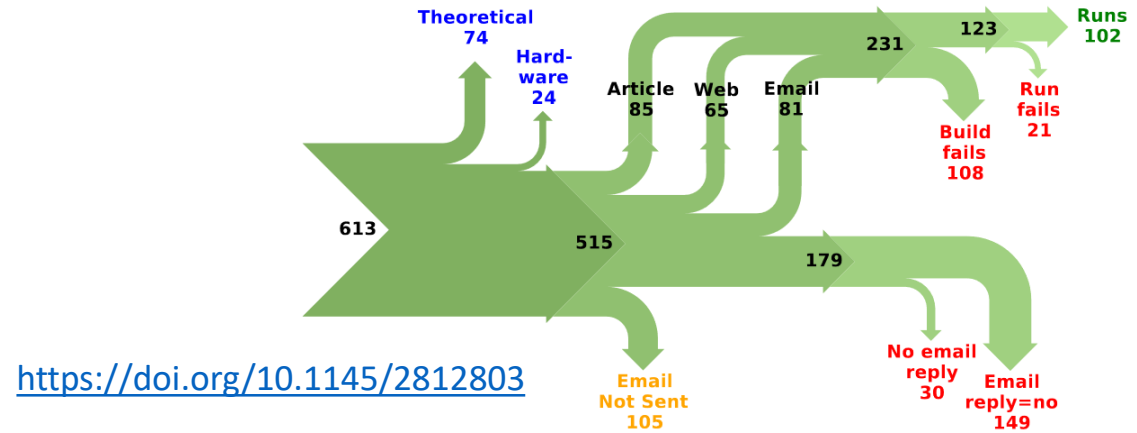
Jupyter Notebook
tools
PDFAnnotator

Diversity is necessary for research, but a challenge for sustainability.

Issues with Software and Reproducibility

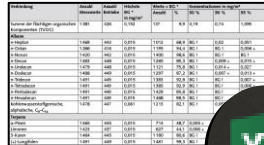
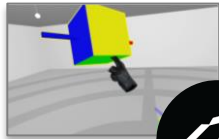
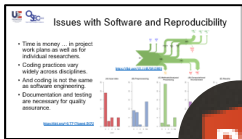
- Time is money ... in project work plans as well as for individual researchers.
- Coding practices vary widely across disciplines.
- And coding is not the same as software engineering.
- Documentation and testing are necessary for quality assurance.

<https://doi.org/10.7717/peerj.5072>



Issues with Tools and Reproducibility

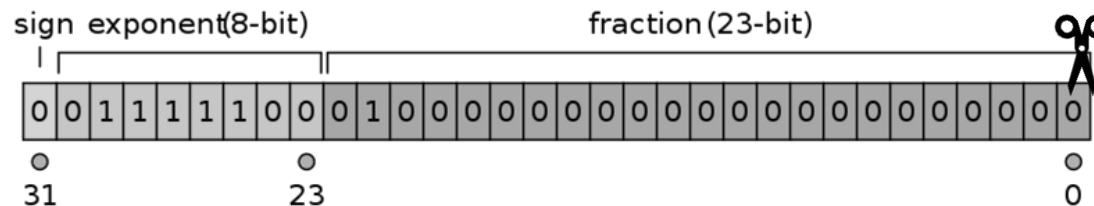
Proprietary data formats are often tied to specific software.



Open data formats can be tied to specific versions of software.

Verbindung	Anzahl Messwerte	Anzahl Beiträge	Höchstes BG in mg/m ³	Werte < BG *	Konzentrationsen in mg/m ³	50 %	90 %	95 %
Summe der flüchtigen organischen Komponenten (TVOC)	1.381	426	0,152	137	9,9	0,19	0,74	1,095
Alkane								
n-Heptan	1.469	443	0,015	1012	69,9	BG.1	0,02	0,051
n-Octan	1.306	414	0,015	1.195	94,8	BG.1	0,06	0,066
n-Nonan	1.420	443	0,015	1.400	98,6	BG.1	0,1	0,1
n-Decan	1.483	448	0,015	1.285	99,9	BG.1	0,099	0,015
n-Undecan	1.479	448	0,015	1.121	79,9	BG.1	0,014	0,027
n-Dodecan	1.486	449	0,015	1.297	97,9	BG.1	0,007	0,013
n-Tridecan	1.491	449	0,015	1.389	92,8	BG.1	0,01	0,007
n-Tetradecan	1.491	449	0,015	1.385	92,8	BG.1	0,006	0,006
n-Pentadecan	1.491	449	0,015	1.429	99,9	BG.1	0,01	0,01
n-Hexadecan	1.491	449	0,015	1.488	99,9	BG.1	0,01	0,01
höchste spezifizierbare, alphabetische, C ₉ -C ₁₄	1.479	447	0,061	1.213	92,1	BG.1	0,007	0,015
Alkene								
n-Propen	1.465	445	0,015	714	48,7	0,009	0,037	0,037
n-Buten	1.423	427	0,015	407	44,1	0,006	0,01	0,01
n-Penten	1.464	445	0,015	1.180	80,8	BG.1	0,01	0,01
n-Hexen	1.491	449	0,015	1.481	99,9	BG.1	0,01	0,01

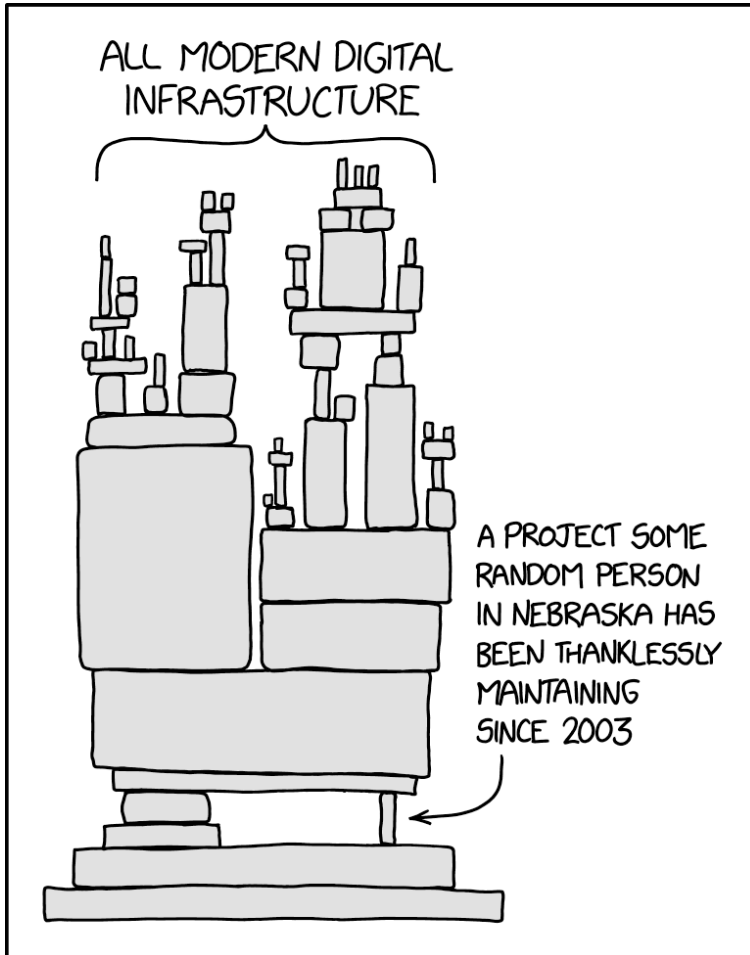
CSV



Software is tied to specific execution environments.

Dealing with data, tools etc. requires dedicated literacy.

Fragile Constructions



<https://xkcd.com/2347/>

Heartbleed

log4j

fMRI bug

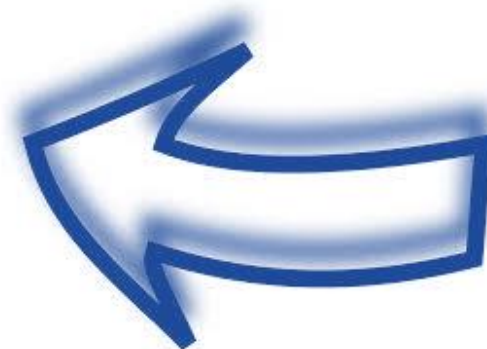
Valuable research funding is wasted
when basic building blocks in software crash!

Solutions – Technical Perspective

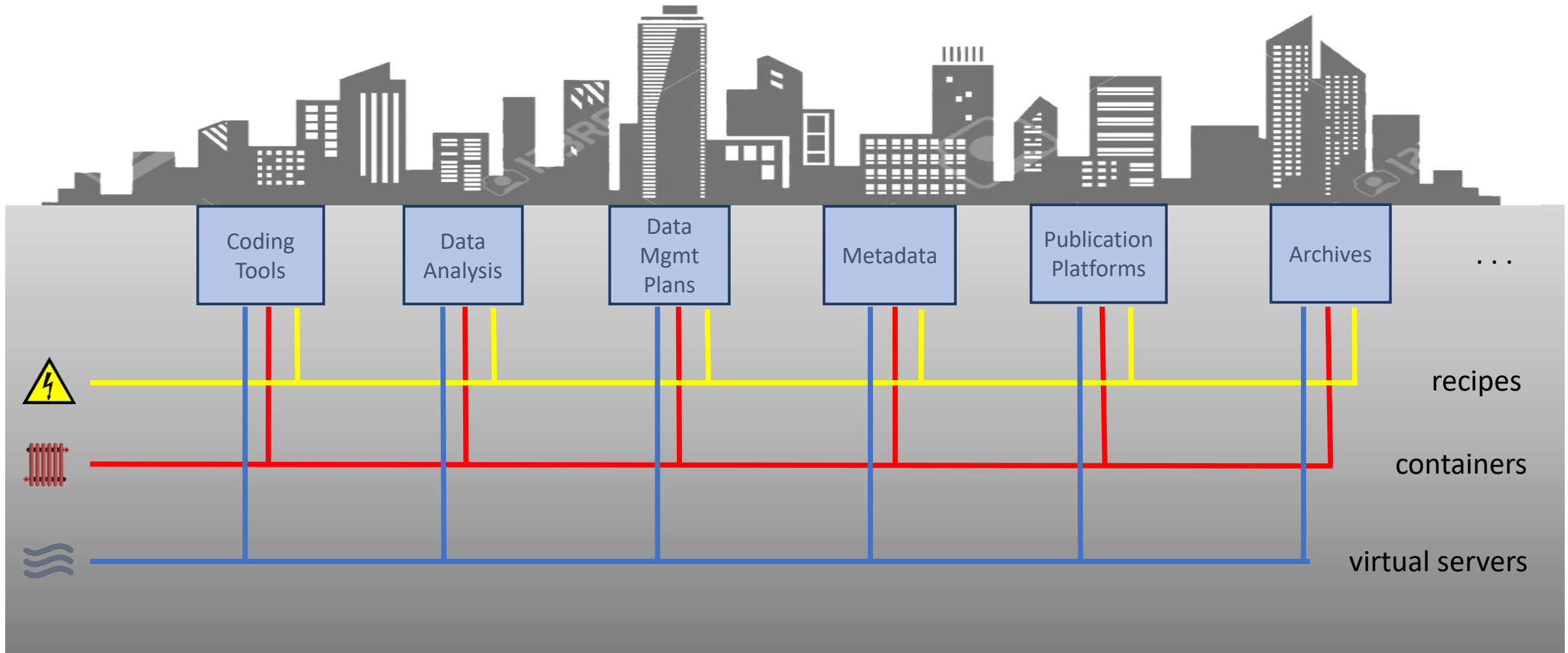
Packaging of

- Data
 - Software
 - Context
- in containers

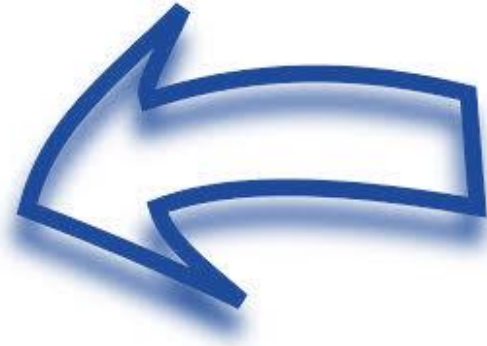
→ Infrastructure for container management



IT Infrastructure



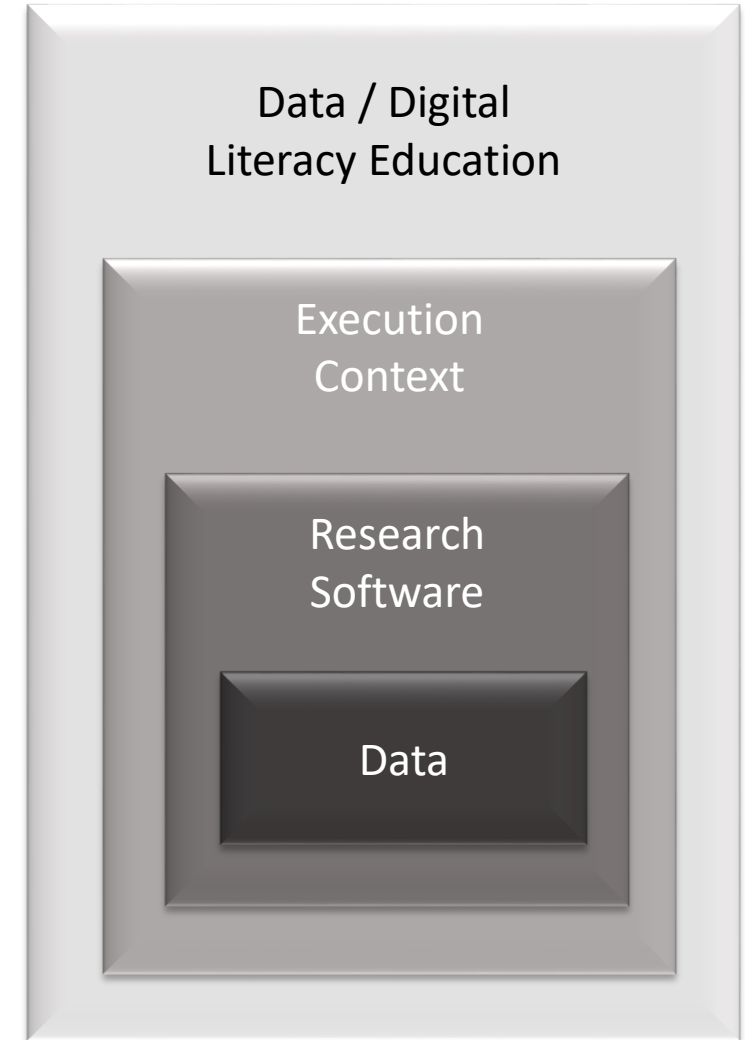
Solutions – Educational Perspective



Dedicated knowledge & skills for

- Data-driven research
- Software engineering
- Computational infrastructure in research

→ Skill development in data & digital literacy



Coding Skills in Academic Curricula

Hackathon «Cultural Data – Data Cultures»


(Digital Humanities)

Dashboard Meine Kurse Dieser Kurs Kurse Hilfe & Support

Blöcke verbergen

Kulturdaten – Datenkulturen

Ein literatur- und kulturwissenschaftlicher Hackathon



Dozent*innen:

Peer Trilcke, Dennis Mischke, Henny Sluyter-Gähje

Die Veranstaltung wird voraussichtlich vollständig virtuell durchgeführt – mit synchronen und asynchronen Anteilen.

Administration

- Kurs-Administration
- Selbst vom abmelden

Navigation

- Dashboard
- Startseite
- Moodle.UP
- Meine Kurse
- SoSe 2021
- Germanist
- KultHac
 - Teilne
 - Badge
 - Komp
 - Bewe
 - Allgen
 - Warmu
 - Datenku
 - Tools
 - Findu
 - Die Ch

Dashboard Meine Kurse Dieser Kurs Kurse Hilfe & Support

Blöcke verbergen

Warm-Up: Kulturdaten in Datenkultureller Bearbeitung

Tools & Ressourcen

Findungsphase

Die Challenges!

Warm-Up: Kulturdaten in Datenkultureller Bearbeitung

Einladungslink Mattermost

DailyArt als Kulturdatenquelle

Ein Projekt von Luise Prager


DailyArt als Kulturdatenquelle

Der Link führt auf mein Padlet, wo ich versucht habe, mei nutzbar.

Spotify - eine globale Kulturda

Ein Projekt von Nadja Jahnke

Spotify - eine globale Kulturdatenquelle?



Three Wishes to the Fairy

1. Recognize the production of research software as a result of research.
(time, funding, skills, ...)
2. Recognize the availability of research software as valuable outcome.
(«publication» counts, metrics, appointment procedures, ...)
3. Recognize the quality of research software as an important issue.
(coding practices, IT infrastructures, support structures, ...)

Recognize
software
as a
first-class
citizen.

May I have Three More?

4. Increase the maturity and integration of current infrastructure & tools.
(Identify and further promote promising approaches.)
5. Adapt the funding instruments to digital infrastructure.
(Add development and maintenance costs to currently predominant human costs.)
6. Establish international collaboration in infrastructure for research software.
(Bring together relevant stakeholders and support their work.)

Promote
the
basic
infrastructure
for
research
software.

There are Plenty of us!



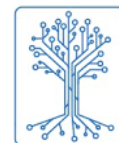
Software Heritage



Software
Sustainability
Institute



nfdi Nationale
Forschungsdaten
Infrastruktur



DEVLOG
RÉSEAU DU DÉVELOPPEMENT LOGICIEL
Recherche et Enseignement Supérieur

Prof. Dr. Ulrike Lucke

University of Potsdam

Institute of Computer Science

ulrike.lucke@uni-potsdam.de